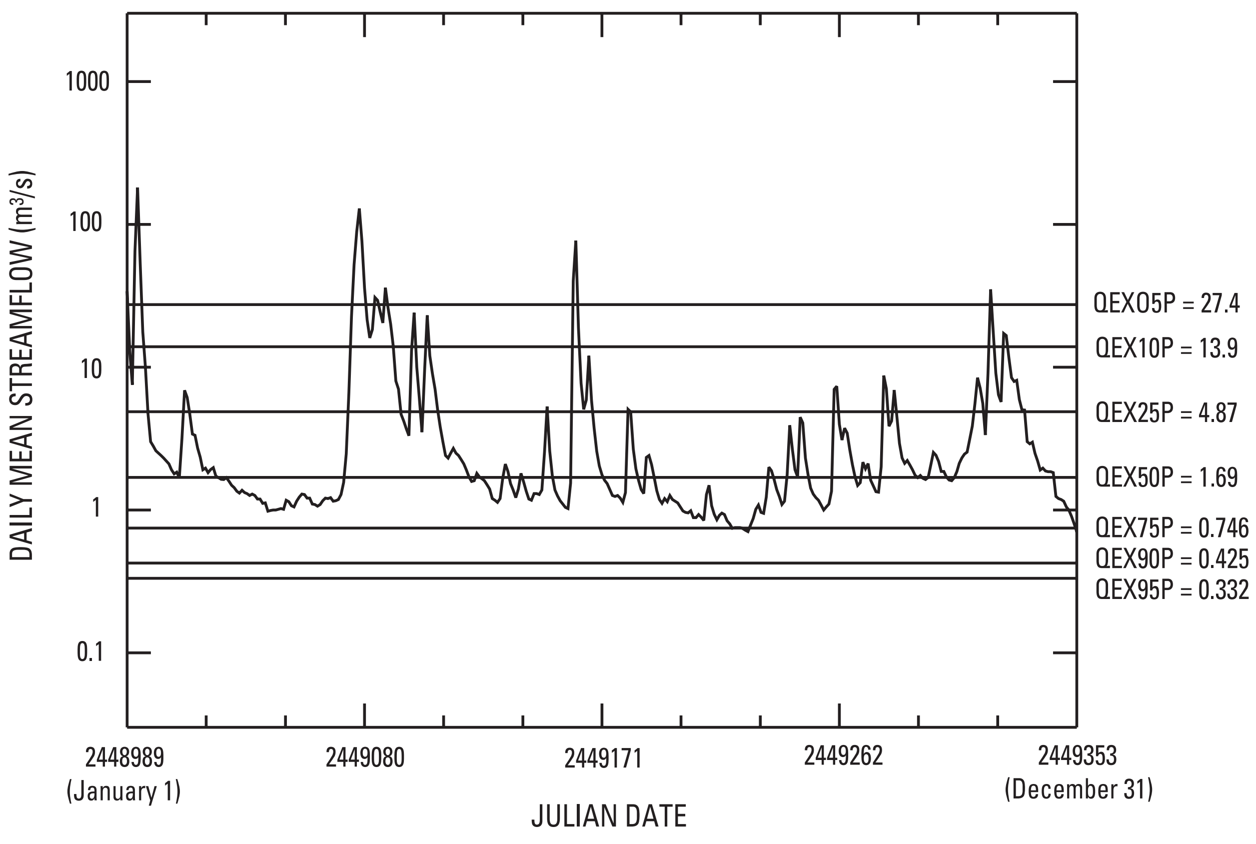
In Cooperation with Environment Canada's National Water Research Institute and the Great Lakes Protection Fund

Calculation of Streamflow Statistics for Ontario and the Great Lakes States



Open File Report 2005-1295

U.S. Department of the Interior
U.S. Geological Survey

Report Documentation Page	*Form Approved* *OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2005	**N/A**	**-**

4. TITLE AND SUBTITLE	
Calculation of Streamflow Statistics for Ontario and the Great Lakes States	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Department of the Interior U.S. Geological Survey 1849 C Street, NW Washington, DC 20240**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**22**	

Cover. Calculated values of the flow-duration statistics (QEX05P, QEX10P, QEX25P, QEX50P, QEX75P, QEX90P, and QEX95P) for the Nith River at New Hamburg relative to daily mean streamflow for 1993. Created by Brian P. Neff and Andrew R. Piggott, July 5, 2005.

Calculation of Streamflow Statistics for Ontario and the Great Lakes States

By Andrew R. Piggott* and Brian P. Neff

* National Water Research Institute, Environment Canada
Burlington, Ontario, Canada

Prepared in Cooperation with Environment Canada's National Water Research Institute and the Great Lakes Protection Fund

Open File Report 2005-1295

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
Gale A. Norton, Secretary

U.S. Geological Survey
P. Patrick Leahy, Acting Director

U.S. Geological Survey, Reston, Virginia: 2005

For sale by U.S. Geological Survey, Information Services
Box 25286, Denver Federal Center
Denver, CO 80225

For more information about the USGS and its products:
Telephone: 1-888-ASK-USGS
World Wide Web: http://www.usgs.gov/

Suggested citation:
Piggott, A.R., Neff, B.P., 2005, Calculation of Streamflow Statistics for Ontario and the Great Lakes States: U.S. Geological Survey Open File Report 2005-1295, 12 p.

CONTENTS

FIGURES

TABLES

Conversion Factors

Inch/Pound to SI

Multiply	By	To obtain
	Length	
inch (in.)	25.4	millimeter (mm)
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
	Area	
square mile (mi^2)	2.590	square kilometer (km^2)
	Flow rate	
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)

Miscellaneous Abbreviations and Glossary

ADAPS – Automated Data Processing System. A database containing USGS streamflow data.

ANNIE - Interactive hydrologic analyses and data management software (Flynn and others, 1995). A USGS program designed to help users interactively store, retrieve, list, plot, check, and update spatial, parametric, and time-series data for hydrologic models and analyses. http://water.usgs.gov/cgi-bin/man_wrdapp?annie

ASCII - American Standard Code for Information Interchange.

DFLOW – Windows-based tool developed by the United States Environmental Protection Agency to estimate user selected design streamflows for low-flow analysis and water quality standards. http://epa.gov/waterscience/dflow/

ESRI – Environmental Systems Research Institute. http://www.esri.com/

FORTRAN – FORmula TRANslation. A high-level programming language for mathematical and scientific purposes.

g77 FORTRAN compiler –A program that translates programs written in FORTRAN into machine language so that they can be executed.

GLI – Graphics Language Interpreter. A graphics system that offers an environment for plotting complex data sets and displaying images. http://iffwww.iff.kfa-juelich.de/gli/

HYDAT – A CD-ROM published by Environment Canada which provides access to HYdro logic DATa contained in the Canadian National Water Data Archive. http://www.wsc.ec.gc.ca/products/hydat/main_e.cfm?cname=hydat_e.cfm

IOWDM – Input and Output for a Watershed Data Management (WDM) file (Lumb and others, 1990). IOWDM is designed to reformat data to and from a WDM file. Data in selected WATSTORE card-image formats and flat file formats can be added to a WDM file using this software. http://water.usgs.gov/software/iowdm.html

LFA – Low Flow Frequency Analysis Package. A hydrologic software package used for the analysis and graphing of the frequency of occurrence of selected low flows, based on streamflow records at one or more gaging sites.
http://www.trentu.ca/wsc/software.shtml

LIBANNE – (Kittle and others, 1989) A collection of related libraries for use in hydrologic applications that provides a WDM data base, graphics, and interactive user interface (AIDE) for hydrologic applications. http://water.usgs.gov/software/libanne.html

PERL – Practical Extraction and Report Language. A general-purpose programming language.
http://www.perl.com/

Red Hat Linux – A computer operating system. http://www.redhat.com/

SWSTAT – Surface Water STATistics (Lumb and others, 1990). A USGS program for developing streamflow statistics, http://water.usgs.gov/software/swstat.html

WATSTORE - The National WATer Data STOrage and REtrieval System. Established in 1972 to provide an effective and efficient means for the processing and maintenance of water data collected through the activities of the U.S. Geological Survey and to facilitate release of the data to the public.
http://ak.water.usgs.gov/Publications/water-data/WY96/watstore.htm

WDM –Watershed Data Management files, identified by the file extension .WDM.

Basic streamflow statistics
CNTVAL – Number of values of streamflow
MAXVAL – Maximum of the values of streamflow
MEANVL – Mean of the values of streamflow
MINVAL – Minimum of the values of streamflow
STDDEV – Standard deviation of the values of streamflow

Flow-duration statistics
QEX05P – 5 percentile flow exceedance
QEX10P – 10 percentile flow exceedance
QEX25P – 25 percentile flow exceedance
QEX50P – 50 percentile flow exceedance
QEX75P – 75 percentile flow exceedance
QEX90P – 90 percentile flow exceedance
QEX95P – 95 percentile flow exceedance

N-day frequency statistics
H01005 – 1-day 5-year recurrence interval high-flow
H01025 – 1-day 25-year recurrence interval high-flow
H01050 – 1-day 50-year recurrence interval high-flow
H01100 – 1-day 100-year recurrence interval high-flow
L07002 – 7-day 2-year recurrence interval low-flow
L07005 – 7-day 5-year recurrence interval low-flow
L07010 – 7-day 10-year recurrence interval low-flow
L07020 – 7-day 20-year recurrence interval low-flow

Calculation of Streamflow Statistics for Ontario and the Great Lakes States

By Andrew R. Piggott and Brian P. Neff

Abstract

Basic, flow-duration, and n-day frequency statistics were calculated for 779 current and historical streamflow gages in Ontario and 3,157 streamflow gages in the Great Lakes states with length-of-record daily mean streamflow data ending on December 31, 2000 and September 30, 2001, respectively. The statistics were determined using the U.S. Geological Survey's SWSTAT and IOWDM, ANNIE, and LIBANNE software and Linux shell and PERL programming that enabled the mass processing of the data and calculation of the statistics. Verification exercises were performed to assess the accuracy of the processing and calculations. The statistics and descriptions, longitudes and latitudes, and drainage areas for each of the streamflow gages are summarized in ASCII text files and ESRI shapefiles.

Introduction

Streamflow statistics calculated using streamflow data summarize the observed flow regime and are frequently cited in water-resource studies. This report describes the calculation of consistent and up-to-date streamflow statistics for streamflow gages in Ontario and the Great Lakes states.

Daily mean streamflow data for all current and decommissioned streamflow gages in Ontario were extracted from the HYDAT CD for 2000 (Environment Canada, 2002). Attributes such as the descriptions, longitudes and latitudes, and drainage areas for the streamflow gages were also extracted from the HYDAT CD. Streamflow data and attributes for 779 streamflow gages in Ontario were retained for processing.

Daily mean streamflow data for the Great Lakes states (Illinois, Indiana, Michigan, Minnesota, New York, Ohio, Pennsylvania, and Wisconsin) were extracted from the U.S. Geological Survey's (USGS) ADAPS database. Attributes for all current and decommissioned streamflow gages in the Great Lakes states were extracted from the USGS's National Water Information System (NWIS) at *http://nwis.waterdata.usgs.gov/usa/nwis/discharge*. Streamflow data for 3,157 streamflow gages in the Great Lakes states were retained for processing; however, attributes analogous to those determined for Canadian streamflow gages were determined for only 3,151 of these streamflow gages. The statistics that were calculated for the remaining six streamflow gages are not included in the final summaries. Length-of-record streamflow data was extracted in both cases. The data for Ontario and the Great Lakes states are for the periods ending on December 31, 2000 and September 30, 2001, respectively.

Figure 1 shows the temporal distribution of the data for Ontario and the Great Lakes states measured in terms of the number of daily mean streamflow values per year. The earliest year of observation is 1860 for Ontario and 1884 for the Great Lakes states. There are 7,202,212 daily mean streamflow values for Ontario and 30,864,754 for the Great Lakes states. The averages of the years of observation weighted relative to the number of values of streamflow-per-year are 1971 and 1966 for Ontario and the Great Lakes states, respectively. Therefore, the temporal distributions of the data are similar despite the large contrast in the total numbers of values.

Spatial distributions of the data measured in terms of the locations of the streamflow gages are shown in figure 2. The distributions of streamflow gages across the Great Lakes states and in southern Ontario are similar, and the reported values of drainage areas have

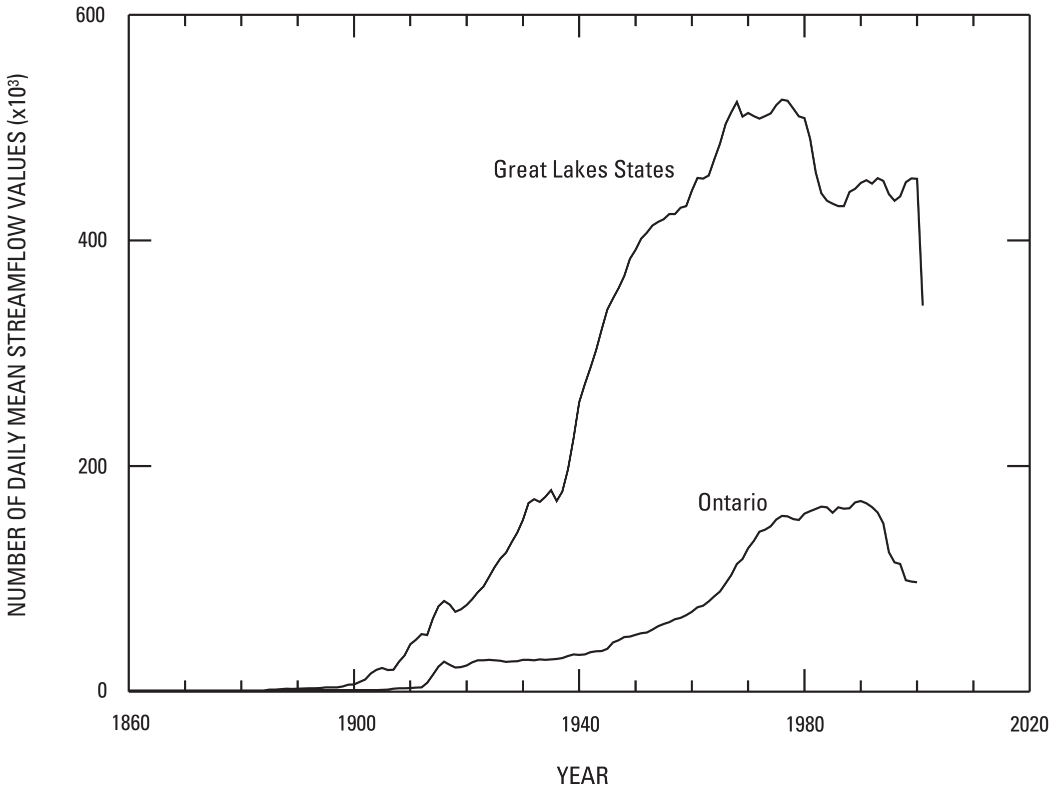


Figure 1. Temporal distributions of the daily mean streamflow data for Ontario and the Great Lakes states measured in terms of the number of values of daily mean streamflow per year.

median values of 316 km^2 and 246 km^2 for Ontario and the Great Lakes states, respectively. Streamflow gages are relatively sparsely distributed and gaged drainage areas are slightly larger in northern Ontario compared to the eight Great Lakes states.

The data for Ontario and the Great Lakes states were extracted from the two databases in different formats. Linux shell and PERL programming were used to organize the data into consistent formats. Julian dates were inserted into the data to facilitate the ordering of the information by date where, for example, the Julian dates 2433376 and 2433377 correspond to the calendar dates April 4 and 5, 1950. The data were partitioned into one ASCII text file per gage where the file name indicates the gage number and where the data are arranged in ascending order relative to the Julian dates. Table 1 lists the first 25 records of the file 02GA018.000, which contains the data for Water Survey of Canada streamflow gage 02GA018, the Nith River at New Hamburg, Ont.. The tab or \t delimited fields of the files are the Julian date and year, month, and day for each observation and the daily mean value of streamflow, in m^3/s. The data were subsequently used in another study to perform hydrograph separation, which required the selection and separation of continuous blocks of daily mean streamflow data. Blocks of data are denoted chronologically in separate files by modifying the 3-digit file extension. For example, there are two discontinuities in the streamflow data for the Nith River at New Hamburg, Ont. and therefore the file 02GA018.000 was partitioned into three files (02GA018.001, 02GA018.002, and 02GA018.003).

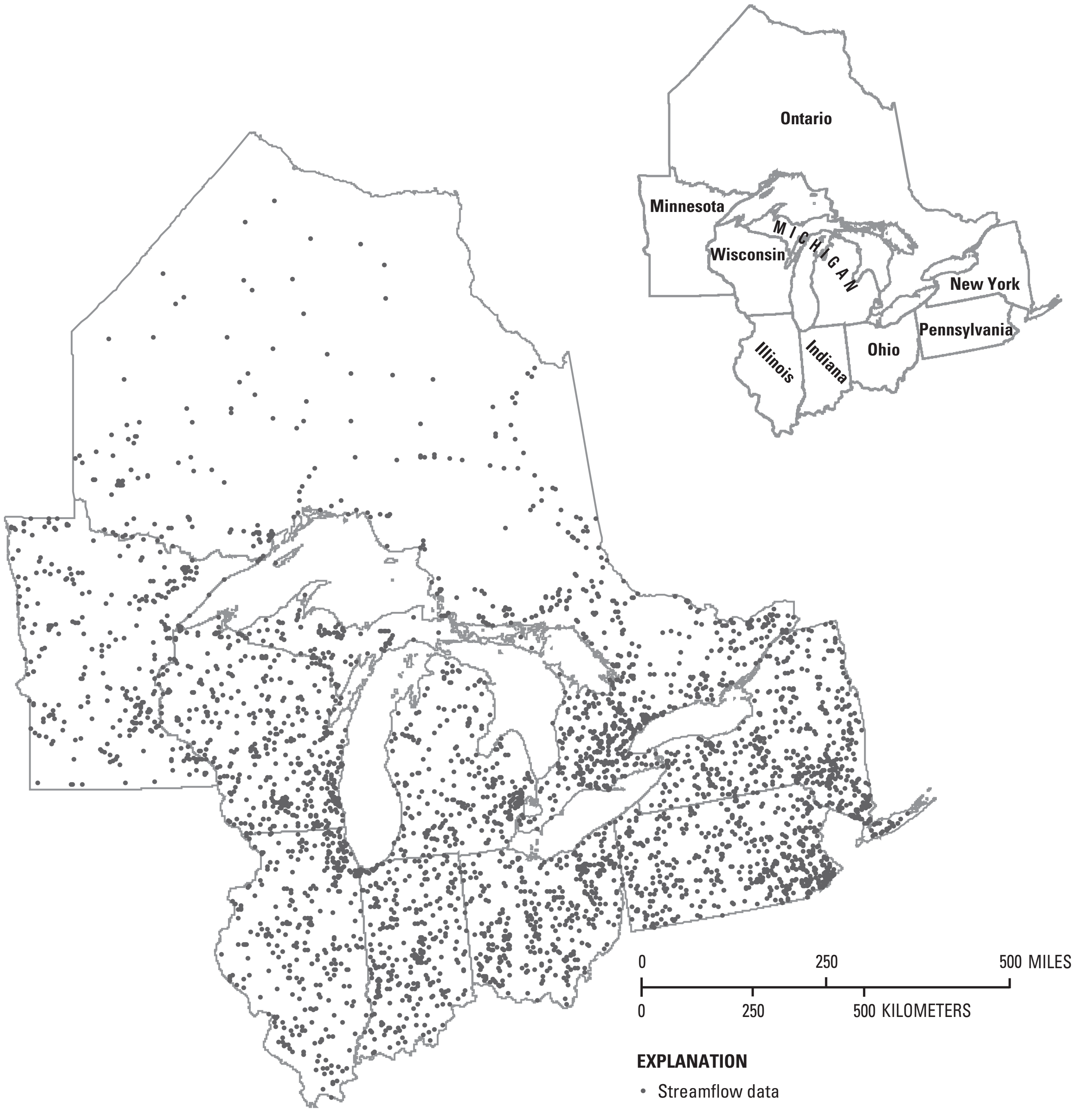


Figure 2. Spatial distribution of the streamflow data.

Table 1. First 25 records of the streamflow gage file 02GA018.000 (Nith River at New Hamburg, Ont.).

[Delimiter, \t, separates Julian date, year, month, day of month, and streamflow in m^3/s.]

Record
2433376\t1950\t4\t4\t192\t
2433377\t1950\t4\t5\t102\t
2433400\t1950\t4\t28\t5.44\t
2433401\t1950\t4\t29\t5.89\t
2433402\t1950\t4\t30\t4.16\t
2433403\t1950\t5\t1\t3.45\t
2433404\t1950\t5\t2\t2.92\t
2433405\t1950\t5\t3\t2.32\t
2433406\t1950\t5\t4\t2.21\t
2433407\t1950\t5\t5\t1.61\t
2433408\t1950\t5\t6\t1.44\t
2433409\t1950\t5\t7\t1.25\t
2433410\t1950\t5\t8\t1.08\t
2433411\t1950\t5\t9\t0.934\t
2433412\t1950\t5\t10\t1.02\t
2433413\t1950\t5\t11\t0.934\t
2433414\t1950\t5\t12\t0.821\t
2433415\t1950\t5\t13\t0.793\t
2433416\t1950\t5\t14\t0.651\t
2433417\t1950\t5\t15\t0.623\t
2433418\t1950\t5\t16\t0.623\t
2433419\t1950\t5\t17\t0.566\t
2433420\t1950\t5\t18\t0.538\t
2433421\t1950\t5\t19\t0.765\t
2433422\t1950\t5\t20\t0.765\t

Calculation of Streamflow Statistics

Three groups of streamflow statistics were calculated. The first group consists of basic statistics and includes the number of daily values of streamflow for each gage and the maximum, mean, minimum, and standard deviation of the daily values (CNTVAL, MAXVAL, MEANVL, MINVAL, and STDDEV, respectively). The second group consists of flow-duration statistics and includes the 5-, 10-, 25-, 50-, 75-, 90-, and 95-percentile flow-exceedance statistics (QEX05P, QEX10P, QEX25P, QEX50P, QEX75P, QEX90P, and QEX95P, respectively). The third group of streamflow statistics consists of n-day frequency statistics and includes the 1-day high-flow statistics with 5-, 25-, 50-, and 100-year recurrence intervals (H01005, H01025, H01050, and H01100, respectively) and the 7-day low-flow statistics with 2-, 5-, 10-, and 20-year recurrence intervals (L07002, L07005, L07010, and L07020, respectively).

The streamflow statistics were calculated using version 4.1 of the USGS SWSTAT software (Lumb and others, 1990), which reads streamflow data from Watershed Data Management (WDM) files and records the calculated values to the WDM files. Version 4.1 of the IOWDM software (Lumb and others, 1990) was used to create WDM files from the ASCII text files, and version 4.1 of the ANNIE software (Flynn and others, 1995) was used to extract the calculated values of the statistics from the WDM files. Version 4.0 of the LIBANNE software (Kittle and others, 1989) was used to compile SWSTAT, IOWDM, and ANNIE, and version 4.5 of the GLI software was used to compile SWSTAT and ANNIE. The statistics were calculated using the Linux operating system and therefore the Unix versions of the SWSTAT and IOWDM, ANNIE, and LIBANNE software were selected. Very few revisions to the source code were required to develop Linux implementations of the SWSTAT and IOWDM, ANNIE, and LIBANNE software using version 7.1 of Red Hat Linux and version 2.96 of the g77 FORTRAN compiler and to process the assembled streamflow data.

Creation of the WDM files and calculation and extraction of the statistics were coordinated using PERL programming where the details of the operations performed by SWSTAT and IOWDM and ANNIE were specified using command files. Tables 2 and 3 list the script (iowdm.pl) and command file (iowdm.log) that were used to translate the ASCII text files (i.e., 02GA018.000) into WDM files (i.e., 02GA018.wdm). Tables 4, 5, and 6 list the script (basic.pl) and command file (basic.log) that were used to calculate the basic statistics and the command file (annie.log) that was used to extract the statistics from the WDM files. Tables 7 and 8 list the script (duration.pl) and command file (duration.template) that were used to calculate the flow-duration statistics where the values of MINVAL and MAXVAL, previously calculated for each gage, were used to derive duration.log from duration.template. Tables 9 and 10 list the script (frequency.pl) and command file (frequency.log) that were used to calculate the n-day frequency statistics.

Table 2. Perl programming script file iowdm.pl, used to translate ASCII text files (i.e., 02GA018.000) into WDM files (i.e., 02GA018.wdm).

Script iowdm.pl

```
#!/usr/bin/perl

while(<STDIN>) {
  chop($_) ;
  $station = $_ ;
  $infile = $station . ".000" ;
  $outfile = $station . ".wdm" ;
  system("cut -f2-6 $infile > iowdm.in") ;
  system("../../iowdm4.1/bin/iowdm <<-EOT
        @../../iowdm.log
        EOT") ;
  system("mv iowdm.out $outfile") ;
}

system("rm ERROR.FIL IOWDM.LOG iowdm.in") ;
```

Table 3. Command file iowdm.log used to translate the ASCII text files (i.e., 02GA018.000) into WDM files (i.e., 02GA018.wdm).

Command File iowdm.log

```
fbiowdm.out
#402
ritsoiowdm.in
#402
fy,f9,m,f9,d,f9,v,f9
#402
t1860
#402
p
#402
rrr
```

Table 4. Perl programming script file basic.pl, used to calculate the streamflow statistics.

Script basic.pl.

```
#!/usr/bin/perl

($outfile) = @ARGV ;

while(<STDIN>) {
  chop($_) ;
  $station = $_ ;
  $infile = $station . ".wdm" ;
  system("cp $infile swstat.in") ;
  system("../../swstat4.1/bin/swstat <<-EOT
        @../../basic.log
        EOT") ;
  system("../../annie4.1/bin/annie <<-EOT
        @../../annie.log
        EOT") ;
  system("tail -n 1 swstat.out |
        tr -s ' ' |
        cut -f7 -d' ' |
        sed -e 's/^/$station CNTVAL /' |
        tr ' ' '\t' >> $outfile") ;
  system("grep 'MINVAL' annie.out |
        sed -e 's/^/$station /' |
        sed -e 's/MINVAL/MINVAL /' |
        tr -s ' ' |
        sed -e 's/ \$//' |
        tr ' ' '\t' >> $outfile") ;
  system("grep 'MAXVAL' annie.out |
        sed -e 's/^/$station /' |
        sed -e 's/MAXVAL/MAXVAL /' |
        tr -s ' ' |
        sed -e 's/ \$//' |
        tr ' ' '\t' >> $outfile") ;
  system("grep 'MEANVL' annie.out |
        sed -e 's/^/$station /' |
        sed -e 's/MEANVL/MEANVL /' |
        tr -s ' ' |
        sed -e 's/ \$//' |
        tr ' ' '\t' >> $outfile") ;
  system("grep 'STDDEV' annie.out |
        sed -e 's/^/$station /' |
        sed -e 's/STDDEV/STDDEV /' |
        tr -s ' ' |
        sed -e 's/ \$//' |
        tr ' ' '\t' >> $outfile") ;
  system("rm swstat.out annie.out") ;
}

system("rm ANNIE.LOG ERROR.FIL
SWSTAT.LOG XPAD.DAT swstat.in") ;
```

Table 5. Command file basic.log, used to calculate the streamflow statistics.

Command File basic.log
foswstat.in
#402
rsbsa1
#402
rd
#213
#213
yes
#213
yes
#402
swstat.out
#402
c
#402
rrr

Table 6. Command file annie.log, used to extract the streamflow statistics from the WDM files.

Command File annie.log
foswstat.in
#402
raeoannie.out
#402
d
#213
no
#213
no
#402
e
#402
Rrr

Table 7. Perl script duration.pl, used to calculate the flow-duration statistics.

Script duration.pl

```
#!/usr/bin/perl

($basic,$outfile) = @ARGV ;

while(<STDIN>) {
  chop($_) ;
  $station = $_ ;
  $infile = $station . “.wdm” ;
  $minval = `grep -w ‘^$station’ $basic |
           grep MINVAL |
           cut -f3` ;
  chop($minval) ;
  if($minval <= 0.0) {
    $minval = 0.001 ;
  }
  if($minval > 1000.0) {
    $minval = 1000.0 ;
  }
  $maxval = `grep -w ‘^$station’ $basic |
           grep MAXVAL |
           cut -f3` ;
  chop($maxval) ;
  system(“sed -e ‘s/MINVAL/$minval/’ ../../duration.template |
         sed -e ‘s/MAXVAL/$maxval/’ > duration.log”) ;
  system(“cp $infile swstat.in”) ;
  system(“../../swstat4.1/bin/swstat <<-EOT
         @./duration.log
         EOT”) ;
  system(“../../annie4.1/bin/annie <<-EOT
         @../../annie.log
         EOT”) ;
  system(“grep ‘QEX95P’ annie.out |
         sed -e ‘s/^/$station /’ |
         sed -e ‘s/QEX95P/QEX95P /’ |
         tr -s ‘ ‘ |
         sed -e ‘s/ \$//’ |
         tr ‘ ‘ ‘\t’ >> $outfile”) ;
  system(“grep ‘QEX90P’ annie.out |
         sed -e ‘s/^/$station /’ |
         sed -e ‘s/QEX90P/QEX90P /’ |
         tr -s ‘ ‘ |
         sed -e ‘s/ \$//’ |
         tr ‘ ‘ ‘\t’ >> $outfile”) ;
  system(“grep ‘QEX75P’ annie.out |
         sed -e ‘s/^/$station /’ |
         sed -e ‘s/QEX75P/QEX75P /’ |
         tr -s ‘ ‘ |
         sed -e ‘s/ \$//’ |
         tr ‘ ‘ ‘\t’ >> $outfile”) ;
```

Table 7. Perl script duration.pl, used to calculate the flow-duration statistics. (continued)

Script duration.pl (continued)

```
  system(“grep ‘QEX50P’ annie.out |
          sed -e ‘s/^/$station /’ |
          sed -e ‘s/QEX50P/QEX50P /’ |
          tr -s ‘ ‘ |
          sed -e ‘s/ \$//’ |
          tr ‘ ‘ ‘\t’ >> $outfile”) ;
  system(“grep ‘QEX25P’ annie.out |
          sed -e ‘s/^/$station /’ |
          sed -e ‘s/QEX25P/QEX25P /’ |
          tr -s ‘ ‘ |
          sed -e ‘s/ \$//’ |
          tr ‘ ‘ ‘\t’ >> $outfile”) ;
  system(“grep ‘QEX10P’ annie.out |
          sed -e ‘s/^/$station /’ |
          sed -e ‘s/QEX10P/QEX10P /’ |
          tr -s ‘ ‘ |
          sed -e ‘s/ \$//’ |
          tr ‘ ‘ ‘\t’ >> $outfile”) ;
  system(“grep ‘QEXO5P’ annie.out |
          sed -e ‘s/^/$station /’ |
          sed -e ‘s/QEXO5P/QEX05P /’ |
          tr -s ‘ ‘ |
          sed -e ‘s/ \$//’ |
          tr ‘ ‘ ‘\t’ >> $outfile”) ;
  system(“rm annie.out”) ;
}

system(“rm ANNIE.LOG ERROR.FIL SWSTAT.LOG
XPAD.DAT duration.log swstat.in”) ;
```

Table 8. Command file duration.template, used to derive duration.log.

Command file duration.template

```
foswstat.in
#402
rsdsa1
#402
rd
#213
#213
none
#213
no
#213
yes
#213
/dev/null
#402
msMINVAL
#213
MAXVAL
#402
p
#213
#213
#213
#213
#213
#213
#213
no
#213
#213
#213
#213
#213
#213
#213
#213
#213
#213
#213
#213
#213
yes
#402
ra
rrr
```

Table 9. Script frequency.pl, used to calculate the n-day frequency statistics.

Script frequency.pl

```
#!/usr/bin/perl

($outfile) = @ARGV ;

while(<STDIN>) {
  chop($_) ;
  $station = $_ ;
  $infile = $station . ".wdm" ;
  system("cp $infile swstat.in") ;
  system("../../swstat4.1/bin/swstat <<-EOT
        @../../frequency.log
        EOT") ;
  system("../../annie4.1/bin/annie <<-EOT
        @../../annie.log
        EOT") ;
  system("grep 'H01005' annie.out |
        sed -e 's/^/$station /' |
        sed -e 's/H01005/H01005 /' |
        tr -s ' ' |
        sed -e 's/ \$//' |
        tr ' ' '\t' >> $outfile") ;
  system("grep 'H01025' annie.out |
        sed -e 's/^/$station /' |
        sed -e 's/H01025/H01025 /' |
        tr -s ' ' |
        sed -e 's/ \$//' |
        tr ' ' '\t' >> $outfile") ;
  system("grep 'H01050' annie.out |
        sed -e 's/^/$station /' |
        sed -e 's/H01050/H01050 /' |
        tr -s ' ' |
        sed -e 's/ \$//' |
        tr ' ' '\t' >> $outfile") ;
  system("grep 'H01100' annie.out |
        sed -e 's/^/$station /' |
        sed -e 's/H01100/H01100 /' |
        tr -s ' ' |
        sed -e 's/ \$//' |
        tr ' ' '\t' >> $outfile")
  system("grep 'L07020' annie.out |
        sed -e 's/^/$station /' |
        sed -e 's/L07020/L07020 /' |
        tr -s ' ' |
        sed -e 's/ \$//' |
        tr ' ' '\t' >> $outfile") ;
  system("grep 'L07010' annie.out |
        sed -e 's/^/$station /' |
        sed -e 's/L07010/L07010 /' |
        tr -s ' ' |
        sed -e 's/ \$//' |
        tr ' ' '\t' >> $outfile") ;
system("grep 'L07005' annie.out |
        sed -e 's/^/$station /' |
        sed -e 's/L07005/L07005 /' |
        tr -s ' ' |
        sed -e 's/ \$//' |
        tr ' ' '\t' >> $outfile") ;
  system("grep 'L07002' annie.out |
        sed -e 's/^/$station /' |
        sed -e 's/L07002/L07002 /' |
        tr -s ' ' |
        sed -e 's/ \$//' |
        tr ' ' '\t' >> $outfile") ;
  system("rm annie.out") ;
}

system("rm ANNIE.LOG ERROR.FIL SWSTAT.LOG
XPAD.DAT swstat.in") ;
```

Table 10. Command file frequency.log, used to calculate the n-day frequency statistics.

Command file frequency.log
foswstat.in
#402
rsnsa1
#402
rmwdm
#213
#213
#213
#213
#213
/dev/null
#402
d
#213
7
#213
0
#213
0
#213
0
#213
0
#213
0
#213
0
#213
0
#213
0
#213
0
#213
0
#213
#213
#213
0
#213
0
#213
0
#213
0
#213
0
#213
0
#213
0
#213
0
#213
0
#402
arfsa501
#213
502
#402
ro/dev/null
#213
no
#402
arrr

Basic statistics were calculated for all of the streamflow gages. Flow-duration statistics were calculated for all streamflow gages, except for four streamflow gages in the Great Lakes states that have atypical data. Calculation of the n-day frequency statistics was limited to streamflow gages with at least five complete high-flow years (October 1 to September 30) and five complete low-flow years (April 1 to March 31) of data. Thus, n-day frequency statistics were calculated for 601 of the 779 streamflow gages in Ontario and 2,230 of the 3,157 streamflow gages in the Great Lakes states.

The statistics were calculated using the published units of the streamflow data; cubic meters per second (m^3/s) for Ontario and cubic feet per second (ft^3/s) for the Great Lakes states. Figure 3 illustrates the calculated values of the basic statistics for the Nith River at New Hamburg, Ont. relative to streamflow for 1993. Only the maximum, mean, and minimum statistics (MAXVAL, MEANVL, and MINVAL) are absolute quantities that can be plotted on a hydrograph and, in this case, MINVAL = 0 and therefore cannot be plotted using the semilogarithmic format of figure 3. Figures 4 and 5 similarly illustrate the calculated values of the flow-duration and n-day frequency statistics, respectively, relative to the streamflow data.

The streamflow gage attributes and basic, flow-duration, and n-day frequency statistics were compiled into a single table for Ontario and a single table for the Great Lakes states. The tables were then output to the ASCII text files (ontario.txt and greatlakesstates.txt), which were used to create ESRI shapefiles for Ontario (ontario.dbf, ontario.sbn, ontario.sbx, ontario.shp, and ontario.shx) and the Great Lakes states (greatlakesstates.dbf, greatlakesstates.sbn, greatlakesstates.sbx, greatlakesstates.shp, and greatlakesstates.shx). These files and the accompanying files readme.txt and md5sum.txt are distributed with this report.

Verification Exercises

The SWSTAT and IOWDM and ANNIE software include series of tests that compare results calculated using an arbitrary implementation of the software to results calculated using a trusted implementation. All of the tests were successfully completed using the Linux implementations of the software that were used to calculate the streamflow statistics. The results of the tests for the implementation of the SWSTAT software were par-

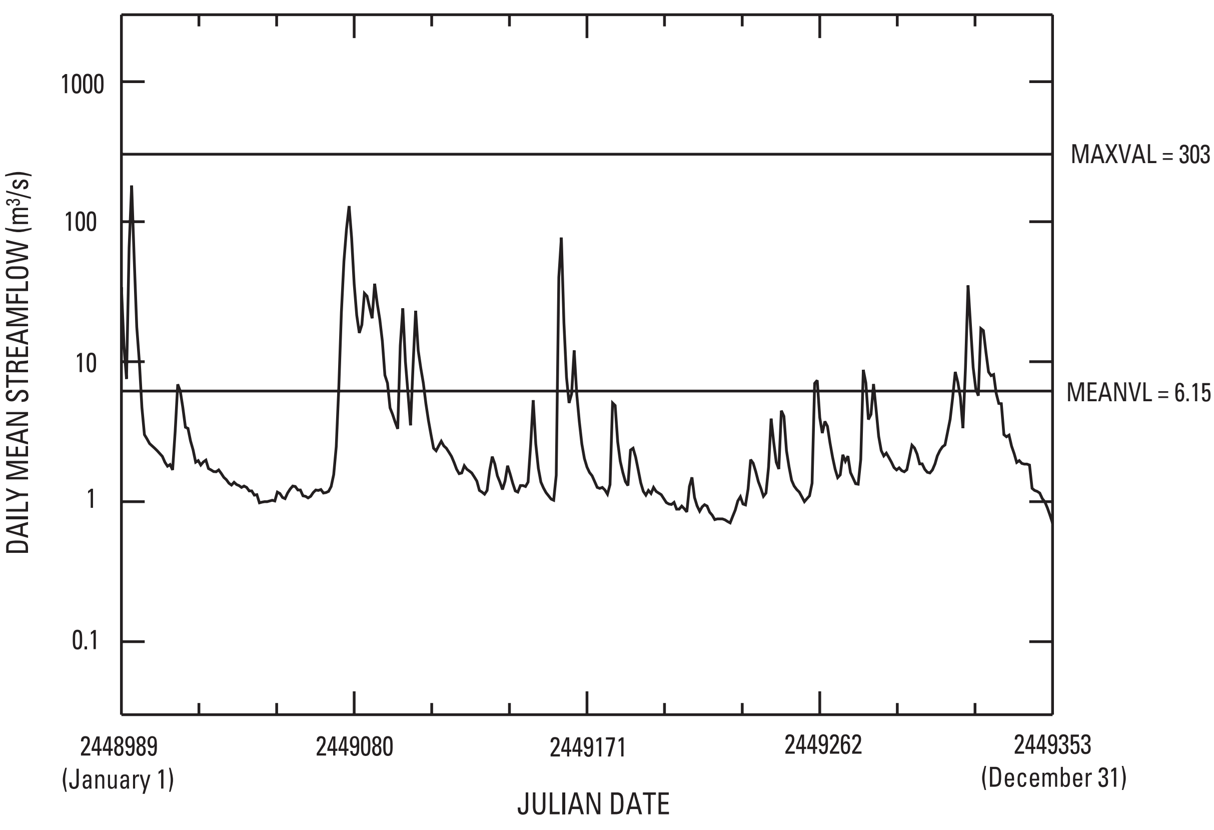


Figure 3. Calculated values of the maximum and mean statistics (MAXVAL and MEANVL) for the Nith River at New Hamburg, Ont. relative to total streamflow for 1993.

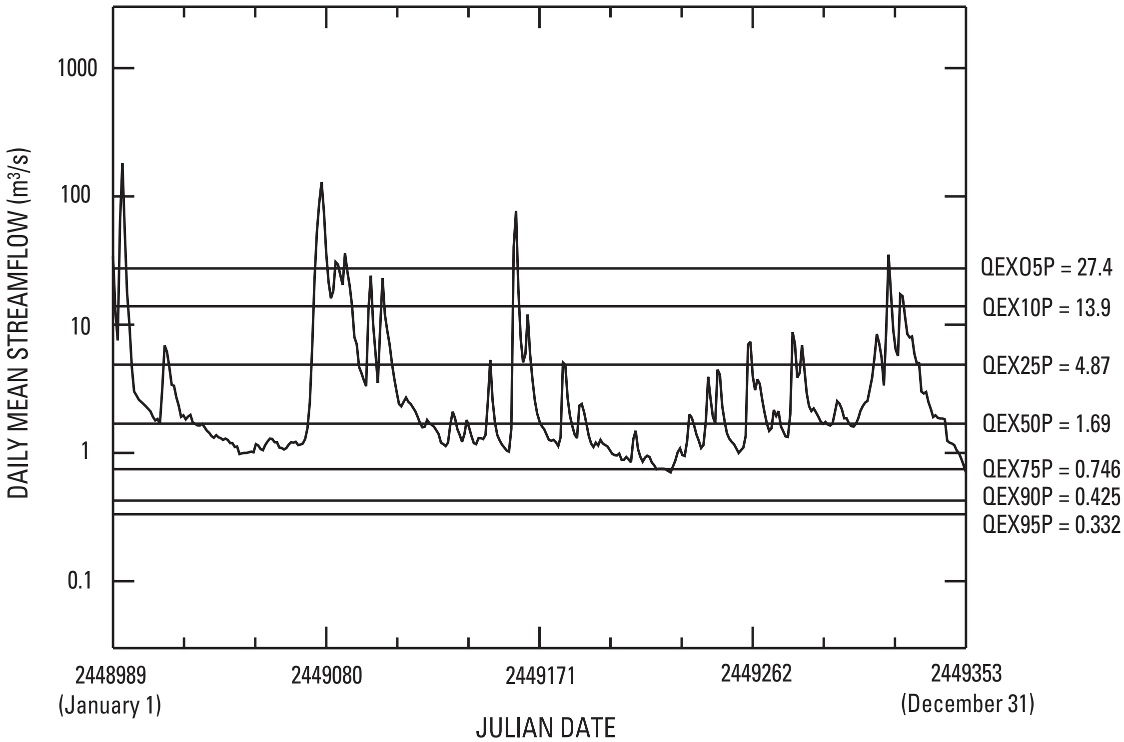


Figure 4. Calculated values of the flow-duration statistics (QEX05P, QEX10P, QEX25P, QEX50P, EX75P, QEX90P, and QEX95P) for the Nith River at New Hamburg, Ont. relative to total stream flow for 1993.

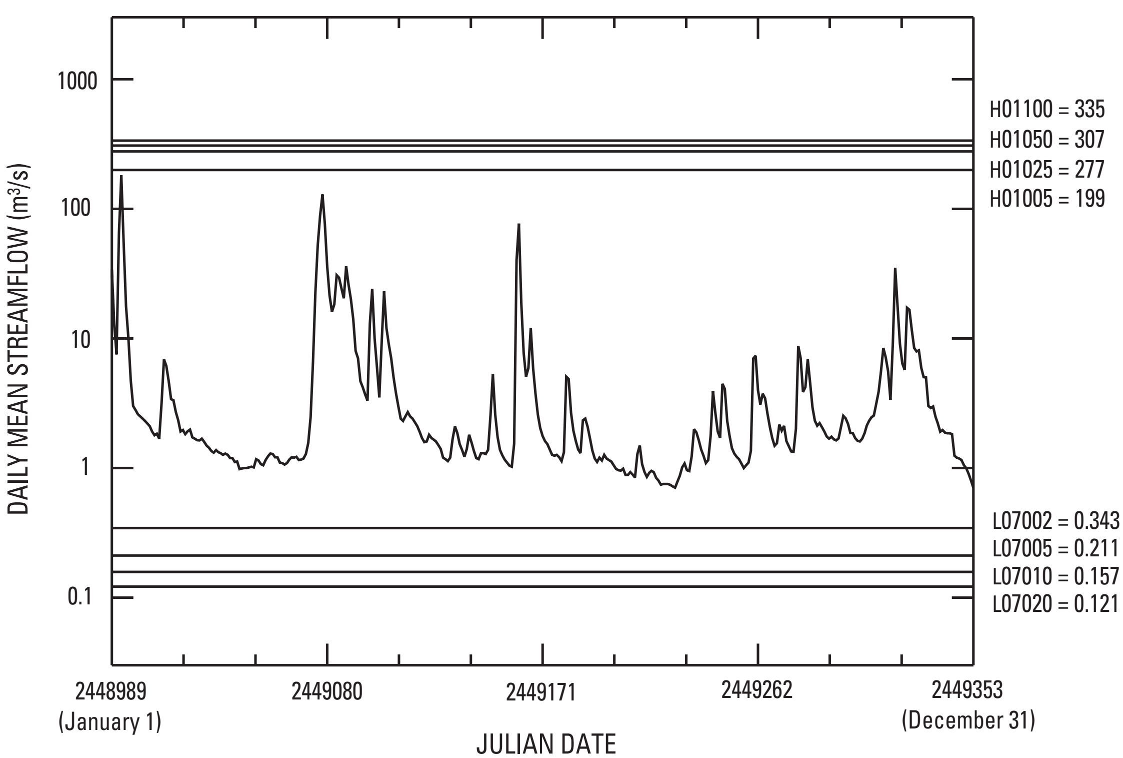


Figure 5. Calculated values of the n-day frequency high-flow (H01005, H01025, H01050, and H01100) and low-flow (L07002, L07005, L07010, and L07020) for the Nith River at New Hamburg, Ont. relative to total stream flow for 1993.

ticularly carefully examined, and it was concluded that the few discrepancies that were observed were due to expected issues such as the differing dates of the calculations and inconsequential issues such as format related differences (e.g., “0.” versus “0.000000”) and computational precision related differences (e.g., “6454.339” versus “6454.338”).

Additional testing was performed by comparing the results calculated using the SWSTAT and IOWDM, ANNIE, and LIBANNE software to results calculated using other procedures and software. The basic statistics for all streamflow gages in Ontario were calculated directly from the ASCII text files and compared to the results calculated using the SWSTAT and IOWDM, ANNIE, and LIBANNE software. The two sets of results matched precisely, which verifies the translation of the ASCII text files into WDM files using iowdm.pl, iowdm.log, and IOWDM; the calculation of the statistics using basic.pl, basic.log, and SWSTAT; and the extraction of the results from the WDM files using annie.log and ANNIE. The flow-duration statistics were also calculated directly from the ASCII text files and compared to the results calculated using the SWSTAT and IOWDM, ANNIE, and LIBANNE software. The two sets of results compare very closely, which verifies the calculation of the statistics using duration.pl, duration.template, and SWSTAT. Finally, the 7-day 10-year recurrence interval low-flow statistic was calculated using the ASCII text files for a random sample of 60 streamflow gages using the U.S. EPA’s DFLOW software (epa.gov/waterscience/dflow/) and compared to the results calculated using the SWSTAT and IOWDM, ANNIE, and LIBANNE software. The two sets of results compare closely, which verifies the calculation process using frequency.pl, frequency.log, and SWSTAT.

In addition, the 7-day 2- and 10-year recurrence interval low-flow statistics calculated using the SWSTAT and IOWDM, ANNIE, and LIBANNE software were compared to results calculated for a selection of streamflow gages using Environment Canada’s LFA software. No significant discrepancy between the two sets of results for either of the two statistics was observed (Richard Pyrce, personal communication, 2005).

The results of these verification exercises appear to indicate that the calculation of the streamflow statistics using the Linux implementations of the SWSTAT and IOWDM, ANNIE, and LIBANNE software and the coordination of these calculations using the scripts and command files listed in tables 2 to 10 are accurate. No verification of the n-day frequency high-flow statistics was performed other than the tests that are included with the SWSTAT software, which appear to have been completed successfully.

Acknowledgments

The authors would like to thank Kathleen Flynn of the U.S. Geological Survey for her assistance in developing the Linux implementations of the SWSTAT and IOWDM, ANNIE, and LIBANNE software. The authors would also like to thank Paul Seelbach of the Michigan Department of Natural Resources and University of Michigan and Ian Cameron of the Ontario Ministry of Natural Resources for their assistance in specifying the streamflow statistics for calculation and Richard Price of Trent University’s Watershed Science Centre for his assistance in verifying the statistics. Calculation of the streamflow statistics was made possible by Environment Canada’s National Water Research Institute and Meteorological Service of Canada, the U.S. Geological Survey, and Great Lakes Protection Fund. Reviews of the manuscript were provided by Denis Healy and Steve Rheaume of the United States Geological Survey. Finally, Lori Fuller of the United States Geological Survey finalized the report layout and all figures.

References Cited

Environment Canada, 2002, HYDAT Version 2000 – 2.01, Surface water and sediment data: Water Survey of Canada, Meteorological Service of Canada, Environment Canada, Toronto, Ontario.

Flynn, K.M., Hummel, P.R., Lumb, A.M., and Kittle, J.L., Jr., 1995, User’s manual for ANNIE, version 2, a computer program for interactive hydrologic data management: U.S. Geological Survey Water-Resources Investigations Report 95-4085, 211 p.

Kittle, J.L., Jr., Hummel, P.R., and Imhoff, J.C., 1989, ANNIE-IDE, A system for developing interactive user interfaces for environmental models (programmers guide): U.S. Environmental Protection Agency, EPA/600/3-89/034, Environmental Research Laboratory, Athens, Ga., 166 p.

Lumb, A.M., Kittle, J.L., Jr., and Flynn, K.M., 1990, Users manual for ANNIE, a computer program for interactive hydrologic analyses and data management: U.S. Geological Survey Water-Resources Investigations Report 89-4080, 236 p. (Documents the 1990 version of IOWDM included in the ANNIE program.)

USGS

Piggott, A.P., Neff, B.P. -- **Calculation of Streamflow Statistics for Ontario and the Great Lakes States** -- Open File Report 2005-1295